

24 NOV 2008 PCT \$

U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371		ATTORNEY'S DOCKET NUMBER: BP100710
INTERNATIONAL APPLICATION NO. PCT/FI98/00441 <i>NOV 24 2008</i> PATENT & TRADEMARK OFFICE		U.S. APPLN. NO. (if known, see 37 CFR 1.5) 09/701069
INTERNATIONAL FILING DATE: 27 May 1998		
PRIORITY DATE CLAIMED: 27 May 1998		
TITLE OF INVENTION: METHOD AND ARRANGEMENT FOR TRANSLATION OF INFORMATION		
APPLICANT(S) FOR DO/EO/US: Ari BECKS and Simo Sakari HEIKKILÄ		
Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:		
1.	<input checked="" type="checkbox"/> This is a FIRST submission of items concerning a filing under 35 U.S.C. 371.	
2.	<input checked="" type="checkbox"/> This is a SECOND or SUBSEQUENT submission of items concerning a filing under 35 U.S.C. 371.	
3.	<input checked="" type="checkbox"/> This express request to begin national examination procedures (35 U.S.C. 371(f)) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371(b) and PCT Articles 22 and 39(1).	
4.	<input checked="" type="checkbox"/> A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.	
5.	<input checked="" type="checkbox"/> A copy of the International Application as filed (35 U.S.C. 371(c)(2)) <ul style="list-style-type: none"> a. <input checked="" type="checkbox"/> is transmitted herewith (required only if not transmitted by the International Bureau). b. <input checked="" type="checkbox"/> has been transmitted by the International Bureau. (see attached copy of PCT/I(B)/308) c. <input type="checkbox"/> is not required, as the application was filed in the United States Receiving Office (RO/US). 	
6.	<input checked="" type="checkbox"/> A translation of the International Application into English (35 U.S.C. 371(c)(2)).	
7.	<input checked="" type="checkbox"/> Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3)). <ul style="list-style-type: none"> a. <input type="checkbox"/> are transmitted herewith (required only if not transmitted by the International Bureau). b. <input type="checkbox"/> have been transmitted by the International Bureau. c. <input type="checkbox"/> have not been made; however, the time limit for making such amendments has NOT expired. d. <input type="checkbox"/> have not been made and will not be made. 	
8.	<input checked="" type="checkbox"/> A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).	
9.	<input type="checkbox"/> An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).	
10.	<input type="checkbox"/> A translation of the annexes of the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).	
Item 11. to 16. below concern document(s) or information included:		
11.	<input type="checkbox"/> An Information Disclosure Statement under 37 CFR 1.97 and 1.98.	
12.	<input type="checkbox"/> An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.	
13.	<input checked="" type="checkbox"/> A FIRST preliminary amendment.	
14.	<input type="checkbox"/> A SECOND or SUBSEQUENT preliminary amendment.	
15.	<input type="checkbox"/> A substitute specification.	
16.	<input checked="" type="checkbox"/> A change of power of attorney and/or address letter.	
	<input checked="" type="checkbox"/> Other items or information:	
Cover Page of the International Publication International Preliminary Examination Report (PCT/IPEA/409) Form PCT/I(B)/308 International Search Report (PCT/ISA/210) Application Data Sheet		

U.S. APPLICATION NO. U. S. PAT. & T. OFFICE 09/701069	INTERNATIONAL APPLICATION NO. PCT/FI98/00441	ATTORNEY'S DOCKET NO. BP100710
		CALCULATIONS PTO USE ONLY
17. <input checked="" type="checkbox"/> The following fees are submitted:		
BASIC NATIONAL FEE (37 CFR 1.492(a)(1)-(5)):		
Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO \$ 1,000.00		
International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO \$ 980.00		
International preliminary examination fee (37 CFR 1.482) not paid to USPTO but international search fee (37 CFR 1.445(a)(2)) paid to USPTO \$ 710.00		
International preliminary examination fee (37 CFR 1.482) paid to USPTO but all claims did not satisfy provisions of PCT Article 33(1)-(4) \$ 690.00		
International preliminary examination fee (37 CFR 1.482) paid to USPTO and all claims satisfied provisions of PCT Article 33(1)-(4) \$ 100.00		
ENTER APPROPRIATE BASIC FEE AMOUNT =		
\$ 1,000.00		
Surcharge of \$130.00 for furnishing the oath or declaration later than 30 months from the earliest claimed priority date (37 CFR 1.492(e)).		
\$ 130.00		
<input checked="" type="checkbox"/> CLAIMS	NUMBER FILED	NUMBER EXTRA
Total claims	15 - 20 -	0
Independent claims	2 - 3 -	0
MULTIPLE DEPENDENT CLAIMS(S) (if applicable)		+ \$270.00
TOTAL OF ABOVE CALCULATIONS =		
\$ 1,130.00		
Reduction of $\frac{1}{2}$ for filing by small entity, if applicable. Applicant claims Small Entity Status under 37 CFR 1.27.		
+ \$ 565.00		
SUBTOTAL =		
\$ 565.00		
Processing fee of \$130 for furnishing the English translation later than months from the earliest claimed priority date (37 CFR 1.49(f)).		
TOTAL NATIONAL FEE =		
\$ 565.00		
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). \$40.00 per property		
+ \$		
TOTAL FEES ENCLOSED =		
\$ 565.00		
		Amount to be refunded:
		charged:
a. <input checked="" type="checkbox"/>	A check in the amount of \$ <u>565.00</u> to cover the above fees is enclosed.	
b. <input type="checkbox"/>	Please charge my Deposit Account No. 25-0120 in the amount of \$ to cover the above fees. A duplicate copy of this sheet is enclosed.	
c. <input checked="" type="checkbox"/>	The Commissioner is hereby authorized to charge any additional fees which may be required by 37 CFR 1.16 and 1.17, or credit any overpayment to Deposit Account No. 25-0120 . A duplicate copy of this sheet is enclosed.	
SEND ALL CORRESPONDENCE TO:		
Customer No. 000466 YOUNG & THOMPSON 745 South 23rd Street 2nd Floor Arlington, VA 22202 (703) 521-2297 facsimile (703) 695-0573		November 24, 2000
By _____		<i>Benoit Castel</i> Benoit Castel Attorney for Applicants Registration No. 35,041

09/701069

534 Rec'd PCT/PTO 24 NOV 2000

PATENTS

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of

Ari BECKS et al.

Serial No. (unknown)

Filed herewith

METHOD AND ARRANGEMENT FOR
TRANSLATION OF INFORMATION

PRELIMINARY AMENDMENT

Commissioner for Patents

Washington, D.C. 20231

Sir:

Prior to the first Official Action and calculation of the filing fee, please amend the above-identified application as follows:

IN THE CLAIMS:

Claim 3, line 1, cancel "or 2".

Claim 4, line 1, change "any of the preceding claims," to --claim 1,--.

Claim 5, line 1, change "any of the preceding claims," to --claim 1,--.

Claim 6, line 1, change "any of the preceding claims," to --claim 1,--.

Claim 7, line 1, change "any of the preceding claims," to --claim 1,--.

Claim 8, line 1, change "any of the preceding claims," to --claim 1,--.

DRAFT - DO NOT CITE

Ari BECKS et al.

Claim 9, line 1, change "any of the preceding claims," to --claim 1,--.

Claim 12, line 1, cancel "or 11".

Claim 14, line 1, change "any of claims 10-13," to --claim 10,--.

Claim 15, line 1, change "any of claims 10-14," to --claim 10,--.

Respectfully submitted,

YOUNG & THOMPSON

By

Benoit Castel

Benoit Castel
Attorney for Applicants
Customer No. 000466
Registration No. 35,041
745 South 23rd Street
Arlington, VA 22202
Telephone: 703/521-2297

November 24, 2000

Method and arrangement for translation of information

5 The invention relates to a method and an arrangement for translating information given as a character string in a first language into a character string in a second language. The invention is advantageously implemented in machine translation of text information.

10 There are previously known methods for linguistically based machine translation of text information. In these methods, the syntax of each language is exactly programmed, so that each language will require a program algorithm of its own. For the storage of vocabularies in different languages, a centralised high-capacity translation memory is used. The EuroTra translation system of the European Union can be mentioned as an example of such a method. Such previously known methods
15 have a number of drawbacks. Exact syntax programming requires most extensive programming operations. Such a syntax algorithm, as well as the necessary translation memory, require a large memory space in the database. Since a translation method operating in this manner is complex, translating within a reasonable time requires an extremely high-powered computer. Due to these
20 shortcomings, the equipment suitable for translation is expensive. Known methods also involve the drawback that updating of the translation algorithm requires programming and updating of the computer program each time.

25 The object of the present invention is to provide a solution for the translation of information which enables the prior art inconveniences described above to be overcome.

30 One idea of the invention is to divide the information to be translated into structural segments and to do the translation by structural segments. The translation is performed on the basis of model segments and rules stored in the knowledge base. The data contained in the knowledge base are advantageously increased so that, in the process of translating, whenever necessary, the user is asked to provide translations of new structural segments over a user interface, these translations being subsequently stored as model segments in the knowledge base. Owing to the
35 solution provided by the invention, the translating equipment requires a smaller memory capacity and a lower processor speed. Moreover, far less programming is required and the operation of the equipment can be developed without program updating.

The method of the invention for machine translation of information given as a character string in a first language into a character string in a second language is characterised by

- 5 - storing model segments in the form of character strings in the first language in the knowledge base and, logically connected to these, model segments in the form of character strings in the second language,
- identifying a structural segment in the character string of said first language following a first rule,
- 10 - comparing said identified structural segment with model segments in the form of character strings in the first language stored according to a second rule,
- striving to select one model segment on the basis of said comparison,
- reading a model segment in the form of a character string in the second language logically connected to the selected model, i.e. equivalent segment and
- 15 - translating said structural segment into said translation segment in the form of a character string in the second language on the basis of said equivalent segment and a third rule.

The arrangement of the invention for translating information given as a character

- 20 string in a first language into a character string in a second language is characterised in comprising
 - knowledge base means for storing model segments in the form of said character strings in the first language and, in logical connection with these, for storing equivalent segments in the form of character strings in the second language, and for
 - 25 storing a first, second and third rule,
 - means for identifying structural segments in said information given as a character string in the first language following a first rule,
 - means for comparing said identified structural segment with the stored model segments in the form of character strings in the first language following a second rule,
 - means for selecting one model segment on the basis of said comparison,
 - means for reading a model segment, i.e. equivalent segment, in the form of a character string in the second language, logically connected to the selected model, in said knowledge base means and
- 30
 - means for translating said structural segment into said translation segment in the form of a character string in the second language on the basis of said equivalent segment and the third rule, said translation segment representing the information to be given in said second language.
- 35

Preferred embodiments of the invention are described in the dependent claims.

5 The invention is described in greater detail below with the aid of the accompanying drawings, of which

figure 1 is a flow chart of a method in accordance with the invention for translating information,

10 figure 2 is a block diagram of an arrangement in accordance with the invention for translating information,

figure 3 illustrates text information divided into structural segments,

15 figure 4 illustrates the translating process of one structural segment with a close model segment appearing in the knowledge base and

figure 5 illustrates the translating process of a structural segment with no close model segment appearing in the knowledge base.

20

Figure 1 illustrates a method in accordance with the invention for translating information. First the information to be translated is read, block 101, and is divided into structural segments according to a first rule, block 102. Subsequently, the first structural segment is read in the untranslated information, block 103. The read

25 structural segment is compared with the model segments stored in the knowledge base, blocks 104 and 110. The comparison is then performed according to a second rule, which determines whether the model segment is close to the structural segment to be translated. If a model segment closely related to this particular structural

30 segment is found in the knowledge base, a model segment i.e. equivalent segment, in the second language logically connected to the close model, block 121, is read in the knowledge base. After this, a translation segment translated into the second language is formed from the structural segment to be translated on the basis of the read equivalent segment following a third rule, block 122. After this it is checked whether there are still untranslated structural segments, block 123. If there are still

35 untranslated structural segments, the process returns to block 103, where the following untranslated structural segment is read for translation. If there are no untranslated structural segments left in block 123, the translation segments are arranged into sentences according to a fourth rule, and the translated information is

then stored. The stored information can be further displayed, e.g. on a screen, or printed out e.g. on paper or a disc, block 124.

- If no model segment close to the structural segment is found in the knowledge base
- 5 in block 110, this particular structural segment is displayed over a user interface means, i.e. a display screen, block 131. The user then feeds the translation of the structural segment, i.e. the equivalent segment, block 132. The structural segment and the equivalent segment are stored for future use as model segments in the knowledge base, blocks 133, 134. After this the process proceeds to block 123 to
- 10 10 continue as explained above. In this case, the equivalent segment is usually directly a translation segment, if the user has been asked to give the translation of the structural segment in the form of the original information. Thus the operation of block 122 is not indispensable in this case.
- 15 15 Said first rule, by which the structural segments are identified, can be based for instance on the identification of "intermediate words" or cases. Intermediate words are for instance prepositions and particles, which usually form standard character strings. Thus, they can be identified by simply comparing the character strings forming each word e.g. with the above known character strings forming an
- 20 20 intermediate word. The identification of cases can be performed e.g. with the aid of suffixes by comparing the last characters of the words with known suffixes. As well known, the character strings forming a word can be separated by means of punctuation. Since a structural segment may advantageously comprise several words, it may also include one or more punctuation marks.
- 25 25 In its most straightforward version, said second rule, by which a structural segment is compared with the model segments, may imply similarity. In this case, exactly the same model segment as the present structural segment to be translated is searched in the knowledge base. However, considering the memory space required for the
- 30 30 knowledge base, it is preferable not to store the different cases of e.g. the model segment separately in the knowledge base, but to identify also a model segment having a different case following the second rule. In this situation, the equivalent segment logically connected to the model segment should also be put in the case needed in order to generate a translation segment. This is done according to the third rule, which consequently covers information about the cases of the language in question.

In many cases, said fourth rule, by which the translation segments are arranged in translated sentences, implies placing the translation segments into the same order in which the structural segments to be translated were in the first language. Yet this order may depend on the language, and hence also said fourth rule is language-specific.

In the storage of the model segments, a type identifier of the model segment can also be advantageously stored. In this case, the type identifier is stored in logical connection with each model segment. If type identifiers are used, various rules can be applied the identification and translation of the structural segment on the basis of the model segment, depending on the type of the structural segment. Types of structural segments are e.g. the object of an action, a proper name, a verb, a place word, an adjective or an idiom. If type identifiers are used, the user is also asked to indicate the type to which the particular structural segment and its translation pertain as the structural segment is translated.

One idea of the invention is to update the knowledge base in the interactively operated translation process. It should be noted that the updating of the knowledge base is not necessarily confined to the storage of new model or equivalent segments, but the rules mentioned above can also be advantageously updated. The updating is then performed e.g. in connection with the translation of a new structural segment fed by the user by identifying the regularity of the input translation.

The translation of one piece of information from a first language into a second language has been described above. The previous updatings of the knowledge base are advantageously utilised in the translation of the subsequent pieces of information. Thus, the process of the invention for translating successive first and second pieces of information may comprise e.g. the following steps:

- reading first information given as a character string in the first language,
- performing the translation of the first information given as a character string in said first language on the basis of data in the knowledge base into first information given as a character string in the second language to the extent this is feasible in terms of the data available in the knowledge base,
- determining the additional data required to complete the translation of the first information given as a character string in the first language into first information given as a character string in the second language,
- feeding said additional data in the knowledge base with a view to update the knowledge base,

- finishing the translation of the first information given as a character string in the first language into first information given as a character string in the second language,
 - storing said first information given as a character string in the second language,
- 5 - storing the second information given as a character string in the first language,
- performing the translation of the second information given as a character string in said first language on the basis of said updated data in the knowledge base into second information given as a character string in the second language.
- 10 Figure 2 is a block diagram of a device arrangement of the invention for the translation of information. The arrangement comprises a disc station 21, a display screen 22 and a keyboard 23 as interface means connected to processor 20. By means of the disc station, information to be translated can be fed from the disc to the device and the translated information can be stored on the disc for use in other devices. The information in question can be transferred between the device and other data processing equipment also over a bus I/O. Display screen 22 can be used to display such structural segments to the user for which no translation is found in the knowledge base. The user can feed the translation of such a structural segment by using keyboard 23. The interface means mentioned above can also be used in the revision and correction of translated information.
- 15 The device shown in figure 2 also comprises an electric memory 24 for temporary storage of structural segments and translation segments, among other things. In addition, the device comprises a mass storage 25 for the storage of the knowledge base, i.e. model segments, type identifiers and rules, as well as programs. For instance a hard disc drive or an optical disc drive can be used as a mass storage. The components mentioned above can be provided by making previously known computer components operate in accordance with the invention using special software. Character strings and other data are advantageously transferred as electric signals between the components.
- 20 The implementation of the invention is by no means confined to the components described above, by contrast, the arrangement of the invention can have many different configurations, which this description enables a person skilled in the art to design.
- 25 Figure 3 illustrates an English sentence divided into structural segments 31, 32, 33 and 34. As shown in the figure, a structural segment typically comprises successive

closely related words in a sentence. Thus a structural segment often includes a punctuation mark separating the words as well.

- Figure 4 illustrates the translation of the first structural segment of the sentence appearing in figure 3 with the aid of one solution of the invention. In the figure, the structural segment 42 to be translated is stored in translation memory 41 and this structural segment is compared with the model segments stored in knowledge base 44. In the case illustrated in figure 4, this particular structural segment has been previously stored in the knowledge base as model segment 45, which is found in the comparison. If, for instance, the present information is to be translated into Finnish, the Finnish model segment 46 logically connected to the English model segment mentioned above is read in the knowledge base. In figure 4 the double line connecting model segments 45 and 46 illustrates a logical connection. When the Finnish model segment has been read it is stored as a translation segment in the translation memory.

- Figure 5 illustrates the translation of the second structural segment shown in figure 3 with the aid of a solution of the invention. In this case, no English structural segment to be translated nor any Finnish equivalent segment has been previously stored as a model segment in the knowledge base. In this case, structural segment 52 to be translated, stored in translation memory 51, is compared with the model segments in the knowledge base, and if the desired equivalent segment is not found in the knowledge base, the structural segment 58 to be translated is shown on the display screen of interface 57. After this, the user feeds the translation 59 of structural segment 58 over the interface in knowledge base 54. In this manner, an English and a Finnish model segment are stored in logical connection in the knowledge base. Then the Finnish translation of the structural segment is stored as a translation segment 53 in translation memory 51.
- Should the structural segments mentioned above reappear in the input information, corresponding model and equivalent segments will be found in the knowledge base, and there will be no need to ask the user to repeat them. If, however, the following input information contains the sentence "we have expanded our operation largely in Finland", "largely" would be a new structural segment. If no close model segment has been previously stored in the knowledge base, the user is asked to give the translation of it and "largely" is stored as a model segment in the knowledge base, and in logical connection with this, also the input translation fed by the user.

- It should be noted that the operation of the equipment can be arranged so that the translation process is first performed by machine for the entire information to the extent allowed by the model segments stored in the knowledge base. After this the user can feed the necessary translations of new structural segments in the knowledge
- 5 base. Such an arrangement has the advantage of the user not having to stay by the computer waiting for the translation process to be completed, but he/she may update the knowledge base with one single input at any suitable moment.
- The model segments can be stored in the knowledge base as pairs of segments, 10 specific pairs of model segments being stored for each language pair. Another way of proceeding is to logically connect model segments in several languages, so that the same model segments can be used as such in the translation of several language pairs. In this case, the model segments of each language can be fed as an input in the knowledge base each time they appear for the first time in the language in question.
- 15 When input information is then fed in the knowledge base during the translation of one language pair, the information contained in the knowledge base will automatically increase also in the other language pairs.
- The solution of the invention is not language-specific on principle, but can be 20 applied to any language pair. Nor is the implementation of the invention restricted to "natural" languages used in ordinary communication, since it can be used to translate any language consisting of character strings into a second language consisting of character strings. Programming languages and data exchange protocols may be mentioned as examples of such other languages.
- 25 The solution of the invention has many advantages over prior art. Its operation requires but little language-specific knowledge for the division of the language into structural segments. A second advantage of the solution consists in additional information being collected in the memory during the process, so that the device
- 30 "learns" new pairs of model segments and rules. Thus, with a straightforward configuration and a small amount of programming and updating it is possible to provide an efficient means for machine translation.
- The solution of the invention is well adapted for use in situations where the 35 arrangement of the invention is used to meet the needs of several users. In this case, the arrangement preferably comprises several interfaces, which may communicate with the knowledge base e.g. over a data transmission network. The knowledge base can then preferably be decentralised in such a way that the first, i.e. the main

knowledge base can be used by a given user group, and the second, i.e. subknowledge base, may be used only by a few in such a given user group. This enables different users to update their own knowledge base e.g. with special vocabularies or expressions, without such knowledge suitable for special purposes

- 5 being used by other users.

In such a decentralised knowledge base, the updating of the first, i.e. the main knowledge base can be performed from the second, i.e. subknowledge bases. Data stored in the second knowledge bases are then transferred to the first knowledge

- 10 base by predetermined criteria. One such criterion may be the incidence of specific data. The data exchange between the knowledge bases can also take place with one common main knowledge database administrator checking and approving each data to be transferred.
- 15 A number of embodiments of the solution in accordance with the invention has been described above. The principle of the invention can, of course, be varied within the scope of protection of the claims, for instance regarding details of the embodiment and fields of application.

Claims

1. A method for machine translation of information given as a character string in a first language into a character string in a second language, comprising
- 5 - storage in the knowledge base of model segments in the form of character strings in said first language, and in logical connection with these, model segments (133, 134) in the form of character strings in the second language,
- identifying a structural segment in the character string of said first language following a first rule (102),
- 10 - comparing said identified structural segment with model segments (104) in the form of character strings in the first language stored according to a second rule,
- striving to select one model segment (110) on the basis of said comparison,
- reading a model, i.e. equivalent segment (121) in the form of a character string in the second language logically connected to the selected model segment, and
- 15 - translating said structural segment into said translation segment in the form of a character string in the second language on the basis of said equivalent segment and a third rule (122),
- characterised** in that the method comprises the identification of an intermediate word and/or a suffix and said first rule is essentially based on the identification of said intermediate word and/or suffix.
- 20
2. A method as claimed in claim 1, **characterised** in that said information to be given as a character string in the second language is generated on the basis of translation segments and a fourth rule (124).
- 25
3. A method as claimed in claim 1 or 2, **characterised** in that, when no model segment to be selected following the second rule is found as a result of the comparison of the structural segments, the structural segment is displayed by means of a user interface (131) and the equivalent segment of the displayed structural segment is stored in the knowledge base by means of the user interface (132, 133).
- 30
4. A method as claimed in any of the preceding claims, **characterised** in that said structural segment comprises a punctuation mark.
- 35
5. A method as claimed in any of the preceding claims, **characterised** in that the type identifier of the model segment is stored in logical connection with the model segment.

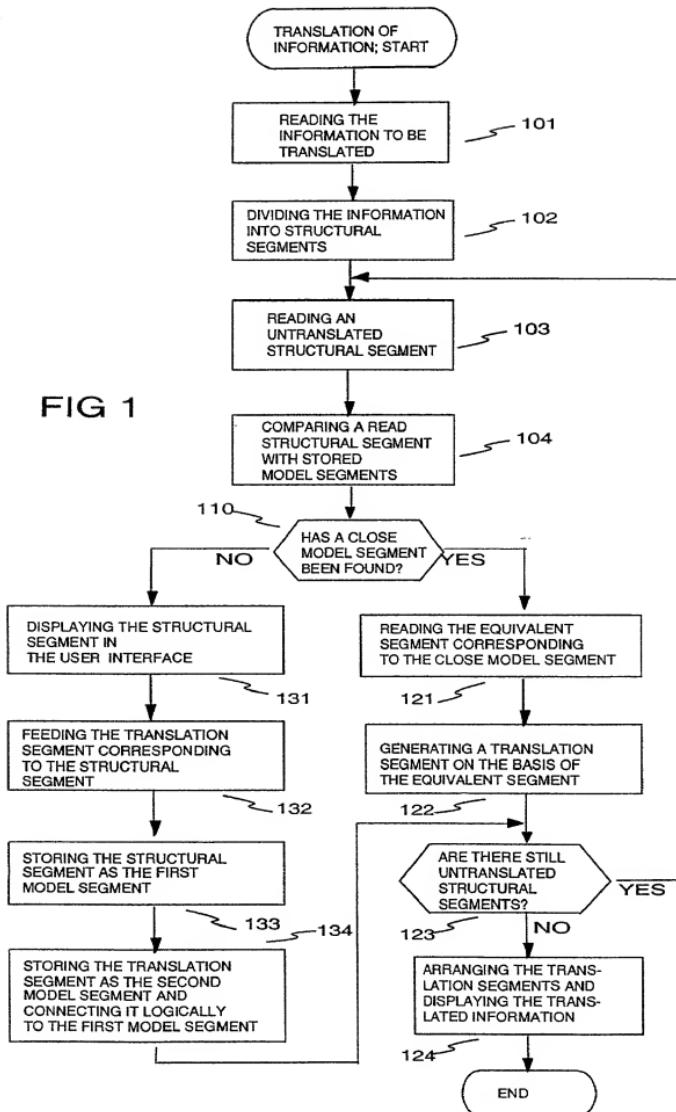
6. A method as claimed in any of the preceding claims, **characterised** in that there are more than two model segments representing different languages logically connected to each other.
- 5 7. A method as claimed in any of the preceding claims, **characterised** in that one of said rules is updated on the basis of output data from the user interface.
8. A method as claimed in any of the preceding claims, **characterised** in that information is fed over the user interface to update the knowledge base with a view
- 10 to translate first information and said input data is used to update other data than those needed for the translation of said first information in said knowledge base.
9. A method as claimed in any of the preceding claims, **characterised** in that it comprises steps for
- 15 - reading the first information given as a character string in the first language,
 - translating the first information given as a character string in said first language on the basis of data in the knowledge base into first information given as a character string in the second language to the extent allowed by the data available in the knowledge base,
- 20 - determining the additional data needed to complete the translation of the first information given as a character string in the first language into first information in the form of a character string in the second language,
 - feeding said additional data in the knowledge base to update the knowledge base,
 - completing the translation of the first information given as a character string in the
- 25 first language into first information given as a character string in the second language,
 - storing said first information given in the second language,
 - reading the second information given as a character string in the first language,
 - translating the second information given as a character string in said first language
- 30 into second information given as a character string in the second language on the basis of said updated data in the knowledge base.
10. An arrangement for translating information given as a character string in a first language into a character string in a second language, comprising
- 35 - knowledge base means (20, 25) for storing model segments in the form of character strings in said first language, and in logical connection with these, equivalent segments in the form of character strings in the second language, and for storing a first, second and third rule,

- means (20, 24) for identifying structural segments in the information given as a character string in said first language following a first rule,
- means (20, 25) for comparing said identified structural segment with the model segments stored in the form of character strings in the first language following a second rule,
- 5 - means (20) for selecting one model segment on the basis of said comparison,
- means (20, 25) for reading the model, i.e. equivalent segment in the form of a character string in the second language logically connected to the selected model segment in said knowledge base means and
- 10 - means (20, 24) for translating said structural segment into said translation segment in the form of a character string in the second language on the basis of said equivalent segment and a third rule, said translation segment representing the information to be given in said second language,
- characterised** in that said means (20, 24) for identifying the structural segment in
- 15 said information given as a character string in the first language comprise means for identifying an intermediate word and/or suffix, said first rule being essentially based on said identification of the intermediate word and/or suffix.
- 20 11. An arrangement as claimed in claim 10, **characterised** in that it further comprises means (20, 25) for generating information to be given as a character string in the second language on the basis of at least two translation segments and a fourth rule.
- 25 12. An arrangement as claimed in claim 10 or 11, **characterised** in that it comprises user interface means (22, 23) for connecting the user to said knowledge base means.
- 30 13. An arrangement as claimed in claim 12, **characterised** in that the user interface means are connected to said knowledge base means over a data transmission network.
- 35 14. An arrangement as claimed in any of claims 10-13, **characterised** in that said knowledge base means comprise a first knowledge base means (25) and a second knowledge base means so that specific users have access to said first knowledge base means and only some of said specific uses have access to said second knowledge base means.
15. An arrangement as claimed in any of claims 10-14, **characterised** in that said knowledge base means comprise a first knowledge base means (25) and a second

knowledge base means, the arrangement comprising means for data input from the user interface means to said second knowledge base means and means for selective transfer of data stored in said second knowledge base to said first knowledge base means.

Abstract

The invention relates to a method and an arrangement for translating information given as a character string in a first language into a character string in a second language. The invention is advantageously implemented in machine translation of text information. One idea of the invention is to divide the information to be translated into structural segments (102) and to perform the translation by structural segments (122). The translation is done on the basis of model segments and rules stored in the knowledge base. The data included in the knowledge base are advantageously increased in such a way that, whenever necessary in the translating process, the user feeds translations (132) of new model segments over the user interface, these translations being subsequently stored as model segments in the knowledge base (133, 134). Owing to the solution of the invention, the translating equipment requires less memory capacity and a lower processor speed. In addition, it requires substantially less programming and the operation of the equipment can be developed without software updating.



09/701069

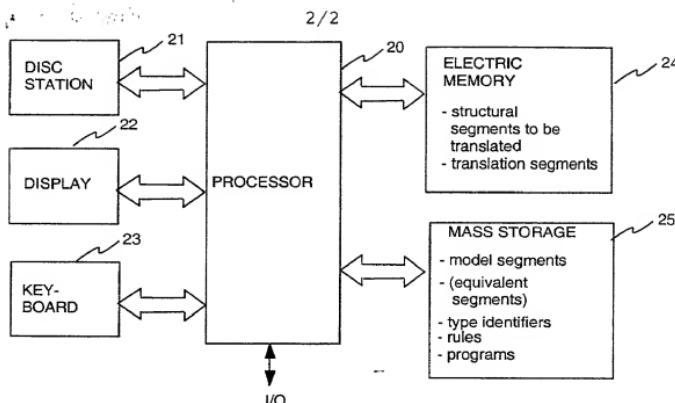


FIG 2

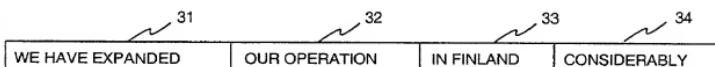


FIG 3

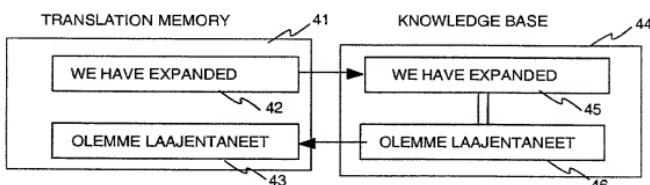


FIG 4

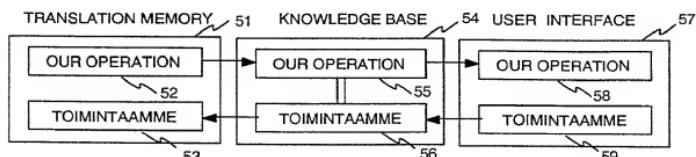


FIG 5

COMBINED DECLARATION AND POWER OF ATTORNEY

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

METHOD AND ARRANGEMENT FOR TRANSLATION OF INFORMATION

the specification of which: *(check one)*

REGULAR OR DESIGN APPLICATION

- is attached hereto.
- was filed on _____ as application Serial No. _____ and was amended on _____ (if applicable).

PCT FILED APPLICATION ENTERING NATIONAL STAGE

was described and claimed in International application No. PCT/FI98/00441 filed on 27 May 1998 and as amended on _____ (if any).

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to patentability as defined in Title 37, Code of Federal Regulations, §1.56.

PRIORITY CLAIM

I hereby claim foreign priority benefits under 35 USC 119 of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application on which priority is claimed.

PRIOR FOREIGN APPLICATION(S)

Country	Application Number	Date of Filing (day, month, year)	Priority Claimed

(Complete this part only if this is a continuing application.)

I hereby claim the benefit under 35 USC 120 of any United States application(s) listed below and, Insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of 35 USC 112, I acknowledge the duty to disclose information which is material to patentability as defined in Title 37 Code of Federal Regulations §1.56 which became available between the filing date of the prior application and the national or PCT international filing date of this application:

(Application Serial No.)

(Filing Date)

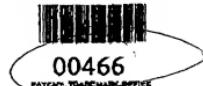
(Status--patented, pending, abandoned)

POWER OF ATTORNEY

The undersigned hereby authorizes the U.S. attorney or agent named herein to accept and follow instructions from Berggren Oy Ab as to any action to be taken in the Patent and Trademark Office regarding this application without direct communication between the U.S. attorney or agent and the undersigned. In the event of a change in the persons from whom instructions may be taken, the U.S. attorney or agent named herein will be so notified by the undersigned.

As a named inventor, I hereby appoint the registered patent attorneys represented by Customer No. 000466 to prosecute this application and transact all business in the Patent and Trademark Office connected therewith, including: Robert J. PATCH, Reg. No. 17,355, Andrew J. PATCH, Reg. No. 32,925, Robert F. HARGEST, Reg. No. 25,590, Benoit CASTEL, Reg. No. 35,041, Eric JENSEN, Reg. No. 37,855, Thomas W. PERKINS, Reg. No. 33,027, and Roland E. LONG, Jr., Reg. No. 41,949,

c/o YOUNG & THOMPSON,
Second Floor,
745 South 23rd Street,
Arlington, Virginia 22202.



Address all telephone calls to Young & Thompson at 703/521-2297. Telefax: 703/685-0573.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Full name of sole or first inventor: Ari BECKS
(given name, family name)

✓

Date 24 November 2000

Inventor's signature *R R*

Citizenship: Finland

Residence: Helsinki, Finland *FI*

Post Office Address: Osuuskunnantie 100 C 3.
FIN-00560 Helsinki, Finland

Full name of second joint inventor, if any: Simo Sakari HEIKKILÄ
(given name, family name)

✓

Inventor's signature *Simo-Sakari Heikkilä* Date 24 November 2000

Citizenship: Finland

Residence: Lahti, Finland *FI*

Post Office Address: Upseerinkatu 1-5 A 2
FIN-19700 Lahti, Finland